神経回路網の数理 : 問題と回答例(補足)(講義ノート)

AUTHOR(S):
田森, 佳秀; 濱中, 雅彦

CITATION:
田森, 佳秀 ...[et al]. 神経回路網の数理 : 問題と回答例(補足)(講義ノート). 物性研究 1989, 52(1): 41-50

ISSUE DATE:
1989-04-20

URL:
http://hdl.handle.net/2433/93612

物性研究52−1(1989−4)

神経回路網の数理−−−問題と回答例(補足)

東北大 工　　田森佳秀　濱中雅彦

(1989年3月4日受理)

はじめに

先に東北大学理学部で行なわれた、甘利先生の大学院講義で、いくつかの小問題が提起された。そのうちで、講義ノートの中で回答の与えられていないものについて、回答を試みたものを、回答例として、甘利先生の講義ノートに添えることにした。この中のいくつかは、学生のリポート提出用として出されたものである。全体としては、寸断された、まとまりのないものになってしまったが、それぞれの問題が独立したものなのでやむをえなかった。

問題1(ボルツマンマシンに関する問題)

ニューロンは、そのニューロンへの入力の総和があるしきい値を越えると興奮する素子であると考えられている。これはi番目のニューロンへの入力の総和をu_i、そのニューロンの状態をx_iとすると、

$$x_i = sgn\ (\ u_i\ ) = \begin{cases} 1 & \text{if } u_i > 0 \\ -1 & \text{if } u_i < 0 \end{cases} \tag{2.1}$$

のように、モデル化することができる(2値モデル)。[1] このように普通の2値モデルでは、u_iによって一意的にx_iが決定されたが、ボルツマンマシン[2] ではそこに確率を導入して、新しい状態を次のような確率で1にする。

$$f\ (\ u_i\ ) = [\ 1 + \exp\ (\ -\ 2\ u_i\ /\ T\ )\ ]^{-1} \tag{2.2}$$

このとき、系がある準安定状態xにある確率が

$$P\ (\ x\ ) = \frac{1}{Z} \exp\ (\ -\ E\ (\ x\ )\ /\ T\ ) \tag{2.3}$$

となることを示せ。

[回答例]　まず、ポテンシャル$E\ (x\ )$と入力の総和u_iは

$$E\ (x\ ) = -\ \frac{1}{2} \sum_{i,j} w_{ij}\ x_i\ x_j \tag{2.4}$$

$$u_i = \sum_{j=1}^{n} w_{ij} x_j \tag{2.5}$$

となる。ここで、1回の状態変化（i 番目のニューロンの状態だけ反転）に対するエネルギー差をみると

$$\begin{aligned}\Delta E &\equiv E(x_i) - E(-x_i) \\ &= -2\sum_{j=1}^{n} w_{ij} x_j = -2 u_i x_i\end{aligned} \tag{2.6}$$

となる。つまり、u_i と x_i が同符号であるようにすると、系はエネルギーの低い方へ動いていくことになる。これより、i 番目のニューロンの状態が $-x_i$ から x_i へ変化することによる遷移確率は、次のようになる。

$$P(x_i \leftarrow -x_i) = [1 + \exp\{(E(x_i) - E(-x_i))/T\}]^{-1} \tag{2.7}$$

時間発展の方程式、

$$\dot{P}(x) = \sum_{i=1}^{n} [P(x_i \leftarrow -x_i)P(-x_i) - P(-x_i \leftarrow x_i)P(x_i)] \tag{2.8}$$

を仮定し、平衡状態での詳細釣合いを考えると、

$$P(x_i \leftarrow -x_i)P(-x_i) - P(-x_i \leftarrow x_i)P(x_i) = 0 \tag{2.9}$$

としてよい。これに(2.7)式を代入すると、

$$\begin{aligned}\frac{P(x_i)}{P(-x_i)} &= \frac{P(x_i \leftarrow -x_i)}{P(-x_i \leftarrow x_i)} \\ &= \frac{1 + \exp\{(E(-x_i) - E(x_i))/T\}}{1 + \exp\{(E(x_i) - E(-x_i))/T\}} \\ &= \frac{\exp\{-E(x_i)/T\}}{\exp\{-E(-x_i)/T\}}\end{aligned} \tag{2.10}$$

となる。したがって、規格化のための定数 Z を用いると、(2.3)式が求まる。

ここで遷移確率(2.2)から準安定状態における状態の確率(2.3)を導くというのは、先ず分配関数を基本に置く統計物理からすると違和感のあることであるが、常に変化する入力を受けている神経系では平衡状態を考えることができず、準安定状態の確率がボルツマン因子である

とは限らないことや、膜電位に応じた発火の確率は一個の神経細胞を用いた実験などから決めることのできるということなどから、神経回路網の統計力学では、どちらかというと遷移確率の方が基本的な量と考えられる。

問題2（状態空間におけるローカルミニマムの数の問題）

ホップフィールド型の、対称（$w_{ij} = w_{ji}$）で、フィードバックを持つ神経網は、エネルギー関数を書くことができて、神経網の状態遷移を、エネルギーの低い方へと系が移り行く物理的描像として、とらえることができる。相互作用の強さw_{ij}が、正規分布している場合の準安定状態の数、すなわちエネルギー関数のローカルミニマムの数を、以下のヒントの方法で求めよ。†

［ヒント］　以下の議論では、すべての入力が1である様なパターン$\boldsymbol{x}^1$を、用いて行くが、これは入力の情報空間から、ある一つのパターンを選んだということで、このことにより一般性を失うような事はない。この入力$\boldsymbol{x}^1$によって発生する膜電位$\boldsymbol{u}$は、

$$
{}^t\boldsymbol{u} = W^t\boldsymbol{x}^1 = \begin{pmatrix} \sum_{i=0}^{n} w_{1i} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=0}^{n} w_{ni} \end{pmatrix} \tag{3.1}
$$

$$
<u_i> = <\sum_{k=0}^{n} w_{ik}> = \sum_{k=0}^{n} <w_{ik}> = 0 \qquad (\quad <w_{ik}> = 0) \tag{3.2}
$$

$$
<u_i u_j> = <\sum_{k=0}^{n} w_{ik} \sum_{l=0}^{n} w_{jl}> = <w_{ij} w_{ji}> = {\sigma_w}^2 \tag{3.3}
$$

$$
<u_i^2> = <\sum_{k=0}^{n} w_{ik} \sum_{l=0}^{n} w_{il}> = <\sum_{k=0}^{n} w_{ik}^2> = n{\sigma_w}^2 \tag{3.4}
$$

これらにより、u_iの分布が分かるので、この分布の第一象限の体積（$u_i > 0$となる確率）が、全状態数に対する準安定状態数の割合である。

† これはスピン系の統計力学を用いる方法で、解かれている。（参考文献[3]）

［回答例］　w_{ij}はランダムな大きさを持ち、平均は(3.2)式より0になり、フィードバックされるときに、その符号しか重要でないから、分散が1の正規分布に従っているとしても一般性を失わない。よって、uは次のn次元正規分布$N(0,\Lambda)$に従う。

$$N(0,\Lambda) \equiv p_0\, e^{-\frac{1}{2}u\Lambda^{-1}\,{}^t u} \tag{3.5}$$

但し、Λは、n行n列のマトリクスで、

$$\Lambda = \begin{bmatrix} n & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & n & 1 & & & & \cdot \\ \cdot & 1 & n & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 & n \end{bmatrix} \tag{3.6}$$

であり、その逆行列は、

$$\Lambda^{-1} = b\begin{bmatrix} a & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & a & 1 & & & & \cdot \\ \cdot & 1 & a & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 & a \end{bmatrix} \tag{3.7}$$

$$a = -2(n-1) \tag{3.7a}$$

$$b = \frac{-1}{(2n-1)(n-1)} \tag{3.7b}$$

となる。ここで、$|\Lambda|$を計算しておくと、

$$|\Lambda| = (n-1)^{(n-1)}(2n-1) \tag{3.8}$$

よって規格化定数p_0は次のように求まる。

$$p_0 = \frac{1}{\sqrt{(2\pi)^n}\sqrt{(n-1)^{(n-1)}(2n-1)}} \tag{3.9}$$

このような分布 (3.5) に従う${}^t u$が、　$u_i > 0$, $(i=1, . . . ,n)$となれば、x^1は、再びx^1となり安定である。この事象の確率は、

$$p = p_0\int_0^{+\infty} \cdots \int_0^{+\infty} \prod_{i=0}^{n} du_i\, e^{-\frac{1}{2}u\Lambda^{-1}\,{}^t u} \tag{3.10}$$

$$p = p_0\prod_{i=0}^{n}\left(\int_{-\infty}^{+\infty} du_i 1(u_i)\right) e^{\frac{1}{2(2n-1)(n-1)}(\sum_{k=0}^{n} u_k)^2 - \frac{1}{2(n-1)}\sum_{k=0}^{n} u_k^2} \tag{3.11}$$

ここで、ハバード変換を用いる。

$$p = p_0\sqrt{\frac{2n-1}{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}(2n-1)}\prod_{i=0}^{n}\left[\int_{-\infty}^{+\infty} 1(u)\ e^{\frac{y}{\sqrt{n-1}}u-\frac{u^2}{2(n-1)}}du\right] \tag{3.12}$$

$$=\sqrt{\frac{(n-1)}{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}(n-1)}\left(\frac{1}{\sqrt{2\pi}}\int_{-y}^{\infty} e^{-\frac{1}{2}z^2}du\right)^n$$

$$=\sqrt{\frac{n-1}{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{1}{2}y^2} e^{nf(y)}dy \tag{3.13}$$

但し、

$$f(y) = \ln\ erf(y) - \frac{1}{2}y^2 \tag{3.14}$$

$f'(y_0)=0$, $f''(y_0)<0$ の解 y_0 により、鞍点法を用いると、

$$p = e^{\frac{1}{2}y_0^2+nf(y_0)}$$

$$=2^{(\log_2 e)(\frac{1}{2}y_0^2 + nf(y_0))} \tag{3.16}$$

準安定状態の数はこれに 2^n をかけたものだから、近似的にこれを求めてみると、

$$p \times 2^n \approx 2^{0.28743n} \tag{3.17}$$

となり、これは田中、エドワーズ等の計算と一致する。

問題３（フィードバックのあるネットワークの記憶容量）

Hopfieldモデルのようなフィードバックを持つネットワークの場合、相関記憶の結合強度は次式のようにおける。[1]

$$w_{ij} = \frac{1}{n}\sum_{\mu=1}^{m} x_i^\mu x_j^\mu \tag{4.1}$$

ここで、n はニューロンの数、m は記憶するパターンの数、 x_i^μ $(\mu=1,\ldots,m)$ は μ 番目の記憶パターンにおける i 番目のニューロンの状態を表わす。また、x_i^μ は確率 $\frac{1}{2}$ で＋１か－１の値をとるものとする。

この場合、n に対して m がある程度小さければそれぞれのパターンを正しく記憶することができる。つまり、(4.1)式によって w_{ij} を決めることによって、それぞれの記憶させた状態が安定状態であるようにすることができる。このとき、正しく記憶することのできる重ね度 $r \equiv \frac{m}{n}$ は

$$r < \frac{1}{2 \ln n} \tag{4.2}$$

であることを示せ。[4,5]

［回答例］　まず、記憶させたパターンの中の1つ $\boldsymbol{x}^1$ が入力されたときの出力 $\boldsymbol{x}'$ をみると、

$$x_i{}' = sgn\left\{\sum_{j=1}^{n} w_{ij}\, x_j^1\right\} = sgn\left\{\frac{1}{n}\sum_{j=1}^{n} x_i^1\, x_j^1\, x_j^1 + \frac{1}{n}\sum_{j=1}^{n}\sum_{\mu=2}^{m} x_i^\mu\, x_j^\mu\, x_j^1\right\}$$

$$= sgn\,(\,A\ x_i^1 + N_i\,) \tag{4.3}$$

$$\begin{cases} A \equiv \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i^1\, x_i^1 \quad = 1 \\ N_i \equiv \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\sum_{\mu=2}^{m} x_i^\mu\, x_j^\mu\, x_j^1 \end{cases} \tag{4.4}$$

となる。ここで、n, m が十分大きいとすると、N_i は正規分布にしたがう（$N_i \sim N\,(\,0,\, r\,)$）としてもよい。すると、

$$Prob\,\{\, x_i{}' \neq x_i^1 \,\} = erf\left[-\frac{A}{\sqrt{r}}\right] \tag{4.5}$$

であるから、

$$A' \equiv \frac{1}{n}\sum_{i=1}^{n} x_i^1\, x_i{}' = \Phi\left[\frac{A}{\sqrt{r}}\right] \tag{4.6}$$

この式は、Kinzel[6] によって求められた式と同じである。ここで、正しく記憶できる条件として例えば $A' > 0.99$ とすると、およそ $r < 0.15$ となり、Hopfieldがシミュレーションで得た結果と一致する。[1] しかし、この A' は1ステップ（すべてのニューロンの状態を1回ずつ変化させた）後の値であり、平衡状態に達したときの値ではない。ただし、正しく記憶できている

かどうかを調べるには A' でも十分である。ここで条件をきびしくして、$A' = 1$ となる条件を求めてみる。そのためには(4.6)式では不十分である。

まず、n 個の要素がすべて変わらない確率 Q は、

$$Q = \prod_{i=1}^{n} Prob\ \{\ x_i{}' = x_i^1\ \} = \left\{ 1 - erf\left(\frac{-1}{\sqrt{r}} \right) \right\}^n \tag{4.7}$$

である。このとき、

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-u} \exp\left(-\frac{x^2}{2} \right) dx \approx \frac{1}{u\sqrt{2\pi}} \exp\left(-\frac{u^2}{2} \right) \tag{4.8}$$

の近似を用い、$m = \dfrac{n}{C_n}$ とおくと(4.7)式は、

$$Q = \left(1 - \frac{1}{\sqrt{2\pi}\sqrt{C_n}}\, e^{-\frac{C_n}{2}} \right)^n \tag{4.9}$$

となる。これが $n \to \infty$ で $Q \to 1$ となればいい。

ここで、

$$I_n \equiv \left(1 - \frac{1}{n\,F_n} \right)^n \tag{4.10}$$

とし、F_n を $n \to \infty$ にするとき最も緩やかに発散する級数とする。すると、

$$\ln I_n = n \ln \left(1 - \frac{1}{n\,F_n} \right) = -n \sum_{k=1}^{\infty} \frac{1}{k} \left(\frac{1}{n\,F_n} \right)^k$$
$$= -\frac{1}{F_n} \left(1 + \frac{1}{2\,n\,F_n} + \cdots \right) \tag{4.11}$$

であるから $\ln I_n$ は最も緩やかに 0 に収束するので、I_n は最も緩やかに 1 に収束する。したがって、(4.9)式と(4.10)式を比較して、定数項を省くことによって、

$$\frac{1}{\sqrt{C_n}}\, e^{-\frac{C_n}{2}} = \frac{1}{n\,F_n} \tag{4.12}$$

とおく。すると Q は $n \to \infty$ でぎりぎり 1 に近づく。このとき、

$$e^{-\frac{C_n}{2}} = \exp\left\{-\ln(\, n\, F_n\,) + \frac{1}{2}\ln C_n\right\} \tag{4.13}$$

$$\begin{aligned} C_n &= 2\ln n + 2\ln F_n - \ln C_n \\ &= 2\ln n + 2\ln F_n - \ln(\, 2\ln n + 2\ln F_n + (\, small\ order\,)\,) \end{aligned} \tag{4.14}$$

であり、F_n は n や $\ln n$ に比べて緩やかに発散するので、 F_n の項は無視できる。したがって

$$C_n < 2\ln n - \ln\ln n \tag{4.15}$$

$$m < \frac{n}{2\ln n - \ln\ln n} \tag{4.16}$$

問題４（神経場の分解能を求める問題）

生物は外界からの入力情報を受けながら、その情報に対応しうるような神経系を自動生成しようとする。甘利先生の神経場モデル[7] では、神経細胞を配列した神経場を考え、外界からの入力信号に対する神経場の反応を見ている。いま外界に k 個の信号 $\boldsymbol{x}^{\mu}$ $(\mu=1,\ldots,k)$ があるとすると、ある入力 $\boldsymbol{x}^{\mu}$ に対しては反応し、他の入力に対しては反応しないような細胞が神経場に自動生成される。このような細胞を認識細胞あるいは表現細胞という。ところが、２つの信号 $\boldsymbol{x}^{\mu}, \boldsymbol{x}^{\nu}$ がある程度近い場合には、これらの入力は同じものであるとみなして、神経場の同じ領域が反応した方がいいとも考えられる。これは、その神経場の入力情報に対する分解能を表わすものである。ある同じ認識細胞を反応させるような入力信号の範囲を、その細胞の受容野という。講義ノートの (8.21) ,(8.22) 式における $\dfrac{c'}{c}x_0{}^2$ が分解能を決める。そこで、$||\boldsymbol{x}^{\mu}|| = 1$、$|\boldsymbol{x}^{\mu}\cdot\boldsymbol{x}^{\nu}| < b$ の場合に、受容野内に１つの信号だけがある条件を求めよ。

［回答例］　まず、R にただ１つの信号 $\boldsymbol{x}^1$ だけしか含まれない場合を考える。このとき、$P(\,R\,) = P^1$ であるので

$$X_R\cdot\boldsymbol{x}^{\nu} = \boldsymbol{x}^1\cdot\boldsymbol{x}^{\nu} \tag{5.1}$$

となる。ところで

$$\begin{cases} X_R\cdot\boldsymbol{x}^1 = \boldsymbol{x}^1\cdot\boldsymbol{x}^1 = 1 & (\,\boldsymbol{x}^1\in R\,) \\ X_R\cdot\boldsymbol{x}^{\nu} = \boldsymbol{x}^1\cdot\boldsymbol{x}^{\nu} < b & (\,\boldsymbol{x}^{\nu}\notin R\,) \end{cases} \tag{5.2}$$

であるから、したがって

$$1 > \frac{c'}{c} x_0^2 > b \tag{5.3}$$

となる。

次に、R に2つ以上の信号が含まれる場合を考える。いま、$x^\mu \in R$ とすると

$$P(R) X_R \cdot x^\mu = \sum_{x^\nu \in R} P^\nu x^\nu \cdot x^\mu = P^\mu x^\mu \cdot x^\mu + \sum_{\substack{\nu \neq \mu \\ x^\nu \notin R}} P^\nu x^\nu x^\mu$$

$$\leq P^\mu - P^\mu b + \sum_{x^\nu \in R} P^\nu b = P^\mu (1 - b) + b P(R) \tag{5.4}$$

となる。R には2個以上の信号が含まれるから、$P^\mu \leq \frac{1}{2} P(R)$ となる $x^\mu \in R$ が必ず存在するはずである。したがって

$$X_R \cdot x^\mu \leq \frac{1}{2}(1 - b) + b = \frac{1}{2}(1 + b) \qquad (x^\mu \in R) \tag{5.5}$$

となる。(5.3)式と(5.5)式より、R 内に1つの信号だけがあるためには、

$$1 > \frac{c'}{c} x_0^2 > \frac{1 + b}{2} \tag{5.6}$$

でなければならない。

References

1. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. USA*, vol. 79, pp. 2554-2558, 1982.

2. D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, "A learnig algorithm for Boltzmann Machines," *Cognitive Sci.*, vol. 9, pp. 147-169, 1985.

3. F. Tanaka and S. F. Edwards, "Analytic theory of the ground state properties of a spin glass : I. Ising spin glass," *J. Phys. F: Metal Phys.*, vol. 10, pp. 2769-2778, 1980.

4. D. J. Amit, H. Gutfreund, and H. Sompolinsky, "Storing infinite numbers of patterns in spin-glass model of neural networks," *Phys. Rev. Lett.*, vol. 55, pp. 1530-1533, 1985.

5. S. Amari and K. Maginu, "Statistical neurodynamics of associative memory," *Neural Networks*, vol. 1, pp. 63-73, 1988.

6. W. Kinzel, "Learning and pattern recognition in spin glass models," *Z. Phys. B*, vol. 60, pp. 205-213, 1985.

7. S. Amari and A. Takeuchi, "Mathematical theory on formation of category detecting nerve cells," *Biol. Cybernetics*, vol. 29, pp. 127-136, 1978.